

GHAJAR EXHIBIT 41

12/13/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

David Esiobu

Page 1

IN THE UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
SAN FRANCISCO DIVISION

RICHARD KADREY, et al.,)	
)	
)	
Plaintiffs,)	Case No.
)	
vs.)	3:23-CV-03417-VC
)	
META PLATFORMS, INC.,)	
)	
)	
Defendant.)	
)	

VIDEOTAPED DEPOSITION OF DAVID ESIOBU
* * * HIGHLY CONFIDENTIAL - ATTORNEYS' EYES ONLY * * *
December 13, 2024
Seattle, Washington

Reported by:
Connie Recob, CCR, RMR, CRR
Washington CCR No. 2631
Oregon CCR No. 15-0436
Utah CCR No. 1133171-7801
Idaho CCR No. SRL-1220

DIGITAL EVIDENCE GROUP
1730 M. Street, NW, Suite 812
Washington, D.C. 20036
(202) 232-0646

12/13/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

David Esiobu

Page 8

1 MR. WEINSTEIN: Just here for the
2 defendant is Mark Weinstein from the Cooley LLP law firm for
3 the defendant and the witness. And with me also is Michelle
4 Woodhouse from Meta legal, also for Meta and for the witness.

5 I apologize, Mr. Young.

6 MR. YOUNG: No problem.

7 BY MR. YOUNG:

8 Q. Mr. Esiobu, are you presently employed?

9 A. Yes.

10 Q. Okay. By Meta Platforms?

11 A. Yes.

12 Q. And what is your present title at Meta?

13 A. Software engineer.

14 Q. Great. And how long have you been in that role at Meta?

15 A. Since August '21, so it would be over three years now.

16 Q. Great. And you understand you are being deposed here in a
17 proceeding involving authors who are suing Meta Platforms
18 alleging copyright infringement of their legal works?

19 A. Uh-huh.

20 MR. WEINSTEIN: Just so -- you have to
21 answer "yes" or "no." You can't go "uh-huh," "huh-uh."

22 THE WITNESS: Sure. Sure.

12/13/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

David Esiobu

Page 22

1 A. Similar to at Citrix. So it was producing code, deploying it
2 and planning product changes, that sort of thing.

3 Q. Okay. And then you left Amazon in May of 2020, correct?

4 A. That's right.

5 Q. Then you went to Microsoft?

6 A. Right, yes.

7 Q. Can you generally describe what your responsibilities were at
8 Microsoft?

9 A. The responsibilities were the same.

10 Q. Okay. Great. And then you left Microsoft in August 2021 to
11 go work for Meta, correct?

12 A. That's right.

13 Q. Okay. And you've been at Meta ever since?

14 A. That's right.

15 Q. Okay. Now, at Meta, would it be fair to say that your
16 responsibilities include the training, testing and
17 development of large language models?

18 MR. WEINSTEIN: Object to form.

19 THE WITNESS: The training, testing and
20 evaluation of large language models.

21 BY MR. YOUNG:

22 Q. I was going to say development.

12/13/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

David Esiobu

Page 23

1 A. Development? Aspects of that, yes. Yeah.

2 Q. Now, you also mention the word "evaluation."

3 Evaluation has a very particular meaning in AI training
4 or AI development, correct?

5 A. Generally, yes.

6 Q. Okay. Now, does evaluation of AI models also fit within your
7 purview or responsibilities at Meta?

8 A. That's -- it's -- that's -- there aren't as rigid boundaries
9 around the specific responsibilities, but I have evaluated
10 models in the past, yes.

11 Q. Okay. Great. Now, Mr. Esiobu, when did you start getting
12 interested in machine learning?

13 A. This would have been sometime in 2019 or 2020, during the
14 end -- towards the end of my time at Amazon.

15 Q. Now, did any of your professional responsibilities either at
16 Amazon or Microsoft involve machine learning?

17 A. Yes.

18 Q. Okay. And just briefly describe what your -- what
19 responsibilities at Amazon may have involved machine
20 learning.

21 A. Yes. At Amazon I worked on their photos product,
22 specifically on the sharing features. And within that, we --

12/13/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

David Esiobu

Page 71

1 A. I am not aware of any policy for that.

2 Q. Okay. Thank you.

3 Do you see in Plaintiffs' 1102 there's a bullet you
4 write called "Copyright Notices and Legal Disclaimers"?

5 A. Uh-huh.

6 Q. Can you read that for me, please?

7 A. It says, "Many documents have copyright notices or legal
8 disclaimers. For example, any similarity dot, dot, dot is
9 coincidental and not intended,' near the beginning or end of
10 the document. The notices are probably not ideal to include,
11 and the disclaimers tend to use really similar boilerplate
12 language which can be repetitive across many documents."

13 Q. Okay. Now, did someone at Meta write a script to remove
14 copyright notices from the beginning and end of documents?

15 A. From what I remember, as part of the processing of LibGen,
16 there was -- there was -- there was a step that it would look
17 at the beginning and end of the document to try to identify
18 and -- and remove the copyright notices.

19 Q. And that was automated, correct?

20 A. Yes, yeah.

21 Q. Do you know who wrote that -- the scripts that automated that
22 process?

12/13/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

David Esiobu

Page 72

1 A. So Nikolay Bashlykov would have been the -- the point of
2 contact for this dataset, and so it was likely him.

3 Q. So do you have an understanding of why copyright notices and
4 legal disclaimers would be removed before the -- during the
5 processing of LibGen?

6 A. So my concern here was kind of the repetitive nature of these
7 kinds of notices. So this is sort of like a form of
8 deduplication to avoid, you know -- to -- to avoid kind of
9 the model, like, regenerating text that it had seen before.

10 Q. Because would you agree with the general proposition that if
11 a model is trained on repetitive text multiple times, it is
12 more likely to regurgitate that -- that portion of text?

13 A. I would say yes. The tests that I run suggests that
14 there's -- there's, like, a strong correlation between
15 repetition and -- in the training data and the rate at which
16 that will be reproduced.

17 Q. And that includes repetition across different files, not just
18 within the same file, correct?

19 A. Right.

20 Q. Okay. So, for example -- and would this also be related to
21 the number of epochs a certain dataset would be run?

22 MR. WEINSTEIN: Object to form.

12/13/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

David Esiobu

Page 73

1 THE WITNESS: I believe it's a similar
2 effect. So either the same text in different documents or
3 the same document multiple times during training would --
4 would have a similar effect.

5 BY MR. YOUNG:

6 Q. And what is epoching?

7 A. Epoching is essentially repeating the training dataset
8 multiple times in order to get -- get the model to extract as
9 much value from -- from it.

10 Q. Okay. And is it true, sir, that it is your testimony that --
11 that more epoching would lead to a higher probability of
12 regurgitating the material contained in that dataset?

13 A. Yeah. The tests -- the tests that I've done suggests that
14 any repetition is kind of correlated with an increase to rate
15 of reproducing the training data.

16 Q. Okay. Thank you.

17 Now, going back to 1102, do you see the section of this
18 document labeled as "Raw Notes"?

19 A. Yes.

20 Q. Okay. Now, is that a script directly underneath that?

21 A. Yes.

22 Q. Okay. And did you write that script?

12/13/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

David Esiobu

Page 74

1 A. I did, yes.

2 Q. Okay. What was the purpose of this script?

3 A. So the dataset was at this path, and it was -- I kind of got
4 tired of copy/pasting the same command to -- to look at
5 samples of the document, so I would -- so this was just a
6 shortened form of being able to just pull samples from
7 that -- that directory.

8 Q. Do you know -- so would it be fair to say that manually
9 sampling LibGen was becoming a very laborious task for you?

10 A. Yeah, I guess I would say.

11 Q. So -- so you wrote a script to make it easier for yourself,
12 right?

13 A. Sure.

14 Q. Yeah. Do you know how many titles were contained in LibGen?

15 A. I don't.

16 Q. Okay. Do you know how many titles were contained in LibGen
17 fiction?

18 A. I don't.

19 Q. Okay. But it was -- it was voluminous enough for you to
20 create a script to sample it, right?

21 A. Yeah. I think -- I think most -- most training datasets
22 are -- are large, yeah.

12/13/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

David Esiobu

Page 75

1 Q. Okay. Thank you.

2 By the way, do you see the Comment 7, please?

3 A. Comment 7 says -- should I read it?

4 Q. Yes, can you please read it. Thank you.

5 A. "Currently we are removing all the copyright paragraphs from
6 beginning and the end of the document. We can expand the
7 list of disclaimers that should be removed."

8 Q. Okay. And Comment 7 generally describes what you were
9 describing Mr. Bashlykov was doing, right?

10 A. Yes.

11 Q. Okay. Thank you.

12 Now, going back to this raw notes section, there is
13 a -- you provide a list of 20 documents, I believe, correct?

14 A. Yes.

15 Q. Okay. And this just generally reports your findings with
16 respect to these 20 documents?

17 A. Yes. I wanted to give a sense of what a -- the kinds of --
18 specific examples of the kinds of things I was seeing during
19 the -- reviewing it.

20 Q. Okay. And these 20 documents were documents identified from
21 the script we just talked about, right?

22 A. Yes.

12/13/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

David Esiobu

Page 76

1 Q. Okay. By the way, just going back to the script, there is
2 a -- there is a file path, right?

3 A. Yes.

4 Q. Okay. And the file path is -- starts with checkpoint and
5 ends with JSONL, right?

6 A. That's right.

7 Q. Okay. And there in the middle is a fiction_EN_deduped,
8 right?

9 A. Yes.

10 Q. Okay. Is that the folder to the entire dataset?

11 A. I -- I don't know off the top of my head. I think so, but I
12 don't know off the top of my head.

13 Q. Okay. And then the next -- the next -- after the next slash
14 is fiction EN processed and then a series of numbers and a
15 chunk.00.JSONL, correct?

16 A. Uh-huh.

17 Q. Okay. That is a subpart of the dataset, correct?

18 A. That's right, yes.

19 Q. Okay. Thank you.

20 By the way, do you recall what -- where the directory
21 is that we can find the script that you wrote?

22 A. So this wasn't in -- I didn't store this. I just kind of ran

12/13/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

David Esiobu

Page 77

1 it in my -- my session at the time, so it was sort of like
2 defined in that session. It wasn't stored in any script
3 anywhere.

4 Q. Okay. So this was just a bespoke script that you wrote to
5 help you with this task?

6 A. Yeah.

7 Q. Okay. Thank you.

8 Can you please turn -- can you please look at the entry
9 for Document 4, please.

10 So can you read the second bullet point for me, please?

11 A. Second bullet point, "Page numbers left in."

12 Q. Oh, no, the second full bullet point.

13 A. Oh, sorry. "Copyright notices at the end."

14 Q. Okay. Do you have an understanding of what you meant by
15 that?

16 A. Yes. I was referring to kind of the repetitive sort of like
17 legal text that -- that I would see across some of the
18 documents.

19 Q. Do you have an -- do you have examples of what some of this
20 text was?

21 A. I don't think that I saved them off anywhere, no.

22 Q. Okay. Was it, for example, the word "copyright"?

12/13/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

David Esiobu

Page 78

1 A. It -- it would -- it would have likely been part of some of
2 the examples.

3 Q. Okay. Are you familiar with the copyright symbol?

4 A. Yes.

5 Q. Okay. Would that be something that you would be looking for
6 when you were looking for copyright notices?

7 A. Not in particular. It -- it was mainly about the -- the, you
8 know, common sentences, common phrases, that sort of thing.

9 Q. Okay. Can you look at doc 5, please?

10 A. Yep.

11 Q. Okay. Do you see the reference in the square -- first square
12 bullet point?

13 A. Uh-huh.

14 Q. Do you know -- can you read that passage for me, please?

15 A. Yes. So it says, "So long Nancy Drew, detective, she
16 thought, making her way out of her cabin and up to the crown
17 deck. Nancy Drew, while she's a cruise passenger is all
18 ready for the sailing party to begin."

19 Q. Are you familiar with Nancy Drew?

20 A. Yes.

21 Q. What is Nancy Drew?

22 A. It's -- it's, like, a series of, like, detective novels,

12/13/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

David Esiobu

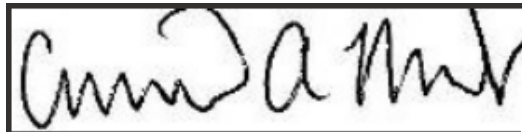
Page 271

1 REPORTER'S CERTIFICATE

2 I, CONNIE A. RECOB, the undersigned Certified Court
3 Reporter, authorized to administer oaths and affirmations in
4 and for the States of Washington, Oregon, Utah and Idaho, do
5 hereby certify that the sworn testimony and/or proceedings, a
6 transcript of which is attached, was given before me at the
7 time and place stated therein; that any and/or all
8 witness(es) were duly sworn to testify to the truth; that the
9 sworn testimony and/or proceedings were by me
10 stenographically recorded and transcribed under my
11 supervision, to the best of my ability; that the foregoing
12 transcript contains a full, true, and accurate record of all
13 the sworn testimony and/or proceedings given and occurring at
14 the time and place stated in the transcript; that a review of
15 which was requested; that I am in no way related to any party
16 to the matter, nor to any counsel, nor do I have any
financial interest in the event of the cause.

17 WITNESS MY HAND and SIGNATURE this 18th day of
December, 2024.

18



19

/s/CONNIE A. RECOB, RMR, CRR

Washington CCR No. 2631

20

Oregon CCR No. 15-0436

21

Utah CCR No. 1133171-7801

22

Idaho CCR No. SRL-1220